

# A Novel Method for Gas Turbine Condition Monitoring Based on KPCA and Analysis of Statistics T2 and SPE

## Authors:

Li Zeng, Wei Long, Yanyan Li

*Date Submitted:* 2019-07-05

*Keywords:* kernel function, SPE statistical model, KPCA, T2 statistical model

## Abstract:

Gas turbines are widely used all over the world, in order to ensure the normal operation of gas turbines, it is necessary to monitor the condition of gas turbine and analyze the tested parameters to find the state information contained in parameters. There is a problem in gas turbine condition monitoring that how to locate the fault accurately if failure occurs. To solve the problem, this paper proposes a method to locate the fault of gas turbine components by evaluating the sensitivity of tested parameters to fault. Firstly, the tested parameters are decomposed by the kernel principal component analysis. Then construct the statistics of T2 and SPE in the principal elements space and residual space, respectively. Furthermore, the thresholds of the statistics must be calculated. The influence of tested parameters on faults is analyzed, and the degree of influence is quantified. The fault location can be realized according to the analysis results. The research results show that the proposed method can realize fault diagnosis and location accurately.

*Record Type:* Published Article

*Submitted To:* LAPSE (Living Archive for Process Systems Engineering)

*Citation (overall record, always the latest version):*

LAPSE:2019.0617

*Citation (this specific file, latest version):*

LAPSE:2019.0617-1

*Citation (this specific file, this version):*

LAPSE:2019.0617-1v1

*DOI of Published Version:* <https://doi.org/10.3390/pr7030124>

*License:* Creative Commons Attribution 4.0 International (CC BY 4.0)

## Article

# A Novel Method for Gas Turbine Condition Monitoring Based on KPCA and Analysis of Statistics $T^2$ and SPE

Li Zeng <sup>1</sup>, Wei Long <sup>2</sup> and Yanyan Li <sup>2,\*</sup><sup>1</sup> School of Aeronautics & Astronautics, Sichuan University, Chengdu 610065, China; zengli\_sichuan@163.com<sup>2</sup> School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065, China; longwei@scu.edu.cn

\* Correspondence: Liyanyan@scu.edu.cn

Received: 24 January 2019; Accepted: 24 February 2019; Published: 27 February 2019



**Abstract:** Gas turbines are widely used all over the world, in order to ensure the normal operation of gas turbines, it is necessary to monitor the condition of gas turbine and analyze the tested parameters to find the state information contained in parameters. There is a problem in gas turbine condition monitoring that how to locate the fault accurately if failure occurs. To solve the problem, this paper proposes a method to locate the fault of gas turbine components by evaluating the sensitivity of tested parameters to fault. Firstly, the tested parameters are decomposed by the kernel principal component analysis. Then construct the statistics of  $T^2$  and SPE in the principal elements space and residual space, respectively. Furthermore, the thresholds of the statistics must be calculated. The influence of tested parameters on faults is analyzed, and the degree of influence is quantified. The fault location can be realized according to the analysis results. The research results show that the proposed method can realize fault diagnosis and location accurately.

**Keywords:** KPCA;  $T^2$  statistical model; SPE statistical model; kernel function

## 1. Introduction

Gas turbines provide power for generators, ships, aircraft, etc. Gas turbines need to withstand the influence of high temperature and high pressure when working. Obviously, the harsh working condition of turbines will definitely lead to the performance degradation of components. Fault occurs when performance degradation is severe. It is essential to locate the fault in time after a fault happens [1–4]. Currently, there are four categories of fault diagnosis—the turbine model-based method, the knowledge-based method, the data-driven-based method, and the techniques fusion-based method [5,6]. The model-based method of the diagnosed object must establish an accurate turbine model and on-line input parameters are employed [5]. Silvio Simani and Farsoni Saverio [7] established an identified fuzzy model which based on the Takagi–Sugeno prototype to detect and isolate the fault. Hector Sanchez and Teresa Escobet [8] established a model and proposed a method to check whether the measurements fall inside the output interval. A diagnosis was proposed based on this model. Method based on knowledge is to essentially formulate the diagnostic problem solving as a pattern recognition problem [9,10]. Zhang, Bingham, and Gallimore [11] proposed two techniques to detect the fault. They promoted the concept of  $y$  indices based on a transposed formulation of data matrix, and residual errors (REs) and faulty sensor identification indices (FSIIs) are introduced in another method. A large number of data must be available if the method based on data-driven is adopted. The potential relationships between these data need to be extracted. Zhu, Ge, and Song [12] proposed a robust variable model driven by the hidden Markov model and a probabilistic model with Student's

t mixture output was designed to tolerate outliers. Furthermore, Zhang Peng [13] studied the Kalman filter and applied it in the location of a fault. He focused on how to establish the linear and nonlinear models of turbine. Based on the models, two faulty location algorithms which apply to the steady working state and dynamic working state respectively were constructed. Vasile Palade, Ron J. Patton, and Faisal J. Uppal [14] applied a neuro-fuzzy technique in an actuator fault location of a gas turbine. Based on learning and adaptation of the TSK fuzzy model, a neuro-fuzzy model was used to generate the residual, and a neuro-fuzzy classifier for the Mamdani model is used to evaluate the residual. Che Changchang, Wang Huawei, and Ni Xiaomei [15] proposed a fault fusion diagnosis model which is based on deep learning. The model analyzes a large number of performance data and obtains fault classification confidence by extracting hidden features from the performance data, then conducts the decision fusion of multiple fault classification results. Tayarani-Bathaie and Khorasani [16] constructed two types of dynamic neural networks to learn the turbine dynamic state. For the measurable variables of the turbine, different neural networks are trained to capture the dynamic relationships. Then, construct a multilayer perception network function to isolate the fault. All model-based methods need to build models that accurately reflect the turbine state. However, due to the large number of turbine parts and the bad working environment, there are too many factors that affect the working state of gas turbine. Thus, it is very difficult to build high-precision models. In addition, the data-driven approaches require sufficient samples to be obtained to locate the fault. Furthermore, the algorithm designer must know the fault generation mechanism and the relationship between these samples. All the above conditions are difficult to meet at the same time.

To avoid the problems mentioned above, and to locate the faults successfully, this paper proposes a fault location approach based on the sensitivity analysis of tested aerodynamic parameters. This approach belongs to the category of data-driven method and the faulty samples are not needed. Firstly, when the turbine is testing, collects the measured data in real-time. Then decomposes the measured data based on the kernel principal component analysis, constructs the Hotelling- $T^2$  ( $T^2$ ) statistic, which is the application of the T-statistic in multivariate analysis in the principal space and squared prediction error (SPE) statistics in residual space after data decomposition. Further, the thresholds of statistics must be calculated, determining whether the fault occurs by comparing the relationship between the  $T^2$  statistic and its threshold. If a fault occurs during detection, we calculate the partial derivatives of the  $T^2$  and SPE statistics to the measured parameters. The greater the values of the partial derivatives, the greater the impacts of the measured parameters to the statistics. According to the working principle of gas turbines, it can be known that the parameters at the outlet of a component will fluctuate firstly and then the fluctuation spreads to other components if a component is faulty. The amplitude of the fluctuation at the outlet of the failed component is the greatest. Obviously, partial derivatives can be used to indicate the degree of influence of the measured parameters when a component fails.

## 2. Materials and Methods

Principal component analysis is a method of data processing which is suitable for linear system and transforms the correlated data into uncorrelated ones by a series of orthogonal changes. Gas turbine is a typical nonlinear system and great error may be caused if PCA is directly used to diagnose the fault of turbine. This paper adopts the kernel principal component analysis to detect the gas turbine fault. By using the kernel function, KPCA has strong nonlinear system processing ability [17,18]. The processes of KPCA are shown below [19–23].

For a given data sample collection,  $x_1, x_2, x_3, \dots, x_q \in \mathbb{R}^n$ , a nonlinear transformation  $\phi$  maps the samples into a higher feature space  $F$ :

$$x \in \mathbb{R}^n \xrightarrow{\phi} \phi(x) \in F \quad (1)$$

where  $\phi(x)$  is the expression of samples in feature space. The covariance of  $\phi(x)$  can be expressed as:

$$C_F = \frac{1}{n} \sum_{i=1}^n \varnothing(x_i) \varnothing(x_i)^T \quad (2)$$

$$\lambda v = C_F v \quad (3)$$

where  $\lambda$  is the eigenvalue of  $C_F$  and  $v$  is the eigenvector of  $C_F$ . Calculating the inner product of  $\varnothing(x_i)$  with  $\lambda v$  and  $C_F v$  respectively:

$$\lambda(\varnothing(x_j) \cdot v) = \varnothing(x_j) \cdot C_F v \quad (4)$$

The eigenvectors can be represented by a series of constants  $\alpha_i$ , as follows:

$$v = \sum_{i=1}^n \alpha_i \varnothing(x_i) \quad (5)$$

Combine Equations (2)–(5):

$$\begin{aligned} \lambda \sum_{i=1}^n \alpha_i [\varnothing(x_j) \cdot \varnothing(x_i)] &= \varnothing(x_j) \cdot \frac{1}{n} \sum_{i=1}^n \varnothing(x_i) \varnothing(x_i)^T \sum_{k=1}^n \alpha_k \varnothing(x_k) \\ &= \frac{1}{n} \sum_{k=1}^n \left[ \left[ \sum_{i=1}^n \varnothing(x_i) \varnothing(x_i)^T \right] [\alpha_k \varnothing(x_j) \cdot \varnothing(x_k)] \right] \end{aligned} \quad (6)$$

Simplify Equation (6) into Equation (7):

$$n\lambda\alpha = K\alpha \quad (7)$$

In Equation (7),  $K = [\varnothing(x_j) \cdot \varnothing(x_i)]$ .  $K$  is a kernel function which calculates the inner product of vectors in high-dimensional feature space. To strengthen the ability of KPCA to deal with nonlinear problem, the Gauss radial basis function is adopted and its expression is:

$$K(x_i, x) = \exp \left[ -\frac{\|x_i - x\|^2}{\sigma^2} \right] \quad (8)$$

Normalize the eigenvectors  $v$  by Equation (8):

$$\langle v_k, v_k \rangle = 1 \quad (9)$$

It can be seen that the vector  $\alpha$  is normalized by Equation (9). Representing the mapped data in the feature space as  $t_k$ , there is:

$$t_k = \langle v_k, \varnothing(x) \rangle = \sum_{i=1}^n \alpha_i^k [\varnothing(x_i) \cdot \varnothing(x)] = \sum_{i=1}^n \alpha_i^k K(x_i, x) \quad (10)$$

where  $\alpha_i^k$  is the  $i$ -th coefficient of the  $k$ -th eigenvalue of matrix  $K$  to eigenvector. The cumulative contribution rate of variance is used to determine the number of principal components which mapped to the feature space. The calculation equation is as follow:

$$\frac{\sum_{k=1}^l \lambda_k}{\sum_{k=1}^n \lambda_k} > \varepsilon \quad (11)$$

where  $l$  is the number of principal component and  $\varepsilon$  is a constant. The value of  $\varepsilon$  reflect the influence of noise. Usually, the value of  $\varepsilon$  is between 0 and 1.

Equations (2)–(11) are the steps to conduct the kernel principal component analysis. To achieve the fault detection of gas turbine components, the statistics of  $T^2$  and SPE must be constructed, as shown below:

$$T^2 = [t_1, t_2, \dots, t_l] \Lambda^2 [t_1, t_2, \dots, t_l]^T \quad (12)$$

where  $\Lambda$  is a diagonal matrix consisting of principal component eigenvalues.  $t_k$  is the mapped data of samples in the feature space.

$$T_{th}^2 = \frac{1(n^2 - 1)}{n(n - 1)} F_{\alpha}(l, n - 1) \quad (13)$$

$T_{th}^2$  is the threshold of the  $T^2$  statistic and  $F_{\alpha}(l, n - 1)$  is upper limit value of F-distribution with confidence level  $\alpha$ :

$$SPE = \|\phi(x) - \phi_k(x)\|^2 = \sum_{i=1}^n t_i^2 - \sum_{i=1}^1 t_i^2 \quad (14)$$

$$SPE_{th} = \theta_1 \left[ \frac{C_a (2\theta_2 h_0^2)^{\frac{1}{2}}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (15)$$

$$\theta_1 = \sum_{i=k+1}^n \lambda_i, \theta_2 = \sum_{i=k+1}^n \lambda_i^2, \theta_3 = \sum_{i=k+1}^n \lambda_i^3, h_0 = 1 - 2\theta_1\theta_2 / (3\theta_2^2) \quad (16)$$

Based on the KPCA introduced from Equations (2)–(15), a fault diagnosis algorithm is designed to determine whether the turbine component is fault. The tested parameters include all the values of total temperature and total pressure at the outlet of gas turbine components. Decomposing these parameters by kernel principal component analysis, construct the  $T^2$  statistic, SPE statistic, and their corresponding thresholds. Determine whether a fault occurs by comparing the relationship of the  $T^2$  statistic and its threshold.

This section focuses on how to locate the fault when the failure occurs. By calculating the partial derivatives of statistics  $T^2$  and SPE to the tested parameters, the sensitivity of tested parameters to the fault can be expressed, and the location of the fault can be determined according to the sensitivity. For the  $T^2$  statistic, the greater the value of sensitivity is, the more likely it is the location of the fault. Kernel function analysis is the most important step in sensitivity calculation, so we make the following changes to the kernel function:

$$K(x_i, x_j) = \exp[-\|x_i - x_j\|^2 / \sigma^2] = K(v \cdot x_i, v \cdot x_j) = \exp[-\|v \cdot x_i - v \cdot x_j\|^2 / \sigma^2] \quad (17)$$

$v = [v_1, v_2, \dots, v_n]$ ,  $v_i = 1$ ,  $n$  is the number of categories of measured parameters,  $x_i$  is the  $i$ -th measurement vector consisting of different measured parameters. Calculating the partial derivative of kernel function to  $v_k$ , there is:

$$\begin{aligned} \frac{\partial K(x_i, x_j)}{\partial v_k} &= \frac{\partial K(v \cdot x_i, v \cdot x_j)}{\partial v_k} = -\frac{1}{\sigma^2} (x_{i,k} - x_{j,k})^2 K(x_i, x_j) \\ &= -\frac{1}{\sigma^2} (v_k \cdot x_{i,k} - v_k \cdot x_{j,k})^2 K(v \cdot x_i, v \cdot x_j) \end{aligned} \quad (18)$$

The value of partial derivative indicates the effect of parameters to kernel function.  $x_{j,k}$ ,  $x_{i,k}$  are the  $k$ -th elements of the  $i$ -th and  $j$ -th measured parameters. The partial derivative of the product between kernels can be expressed as:

$$\frac{\partial K(x_i, x_{new}) K(x_j, x_{new})}{\partial v_k} = -\frac{1}{\sigma^2} \left[ (x_{i,k} - x_{new,k})^2 + (x_{j,k} - x_{new,k})^2 \right] \times K(x_i, x_{new}) K(x_j, x_{new}) \quad (19)$$

$x_{new}$  is a vector consisting of measured parameters. Define the partial derivatives of statistics as  $C_{T^2, i, new}$  and  $C_{SPE, i, new}$ , there are:

$$C_{T^2, i, new} = \left| \frac{\partial T_{new}^2}{\partial v_k} \right|, C_{SPE, i, new} = \left| \frac{\partial SPE_i}{\partial v_k} \right| \quad (20)$$

The values of  $C_{T^2,i,new}$  and  $C_{SPE,i,new}$  indicate the sensitivity level of the  $i$ -th element of the statistics. Steps of calculate the  $C_{T^2,i,new}$  are as follows:

$$T^2 = [t_1, t_2, \dots, t_k] \Lambda^{-1} [t_1, t_2, \dots, t_k]^T = \bar{K}_{new}^T \alpha \Lambda^{-1} \alpha^T \bar{K}_{new} = \text{tr}(\alpha^T \alpha^T \bar{K}_{new} \bar{K}_{new}^T \alpha \Lambda^{-1}) \quad (21)$$

$$\bar{K}_{new} = \begin{matrix} K(x_1, x_{new}) - \frac{1}{n} \sum_{j=1}^n K(x_1, x_j) - \frac{1}{n} \sum_{j=1}^n K(x_{new}, x_j) + \frac{1}{n^2} \sum_{j=1}^n \sum_{j=1}^n K(x_j, x_j) \\ K(x_2, x_{new}) - \frac{1}{n} \sum_{j=1}^n K(x_2, x_j) - \frac{1}{n} \sum_{j=1}^n K(x_{new}, x_j) + \frac{1}{n^2} \sum_{j=1}^n \sum_{j=1}^n K(x_j, x_j) \\ \dots \\ K(x_n, x_{new}) - \frac{1}{n} \sum_{j=1}^n K(x_2, x_j) - \frac{1}{n} \sum_{j=1}^n K(x_{new}, x_j) + \frac{1}{n^2} \sum_{j=1}^n \sum_{j=1}^n K(x_j, x_j) \end{matrix} \quad (22)$$

There is:

$$C_{T^2,i,new} = \left| \frac{\partial T_{new}^2}{\partial v_k} \right| = \left| \frac{\partial (\text{tr}(\alpha^T \bar{K}_{new} \bar{K}_{new}^T \alpha \Lambda^{-1}))}{\partial v_k} \right| = \left| \text{tr}[\alpha^T (\frac{\partial \bar{K}_{new} \bar{K}_{new}^T}{\partial v_k}) \alpha \Lambda^{-1}] \right| \quad (23)$$

The calculation steps of SPE are as follows:

$$SPE = K(x_{new}, x_{new}) - \frac{2}{n} \sum_{i=1}^n K(x_i, x_{new}) + \frac{1}{n^2} \sum_{i=1}^n \sum_{i=1}^n K(x_i, x_i) - t_{new}^T t_{new} \quad (24)$$

$$\begin{aligned} C_{SPE,i,new} &= \left| \frac{\partial SPE}{\partial v_i} \right| = \left| -\frac{1}{\sigma^2} \left[ -\frac{2}{n} \frac{\partial \sum_{i=1}^n k(x_i, x_{new})}{\partial v_i} - \frac{\partial t_{new}^T t_{new}}{\partial v_i} \right] \right| \\ &= \left| \frac{1}{\sigma^2} \left[ \frac{2}{n} \frac{\partial \sum_{i=1}^n k(x_i, x_{new})}{\partial v_i} + \frac{\partial t_{new}^T t_{new}}{\partial v_i} \right] \right| \\ &= \left| \frac{1}{\sigma^2} \left[ \frac{2}{n} \frac{\partial \sum_{i=1}^n k(x_i, x_{new})}{\partial v_i} + \text{tr}(\alpha^T \frac{\partial t_{new}^T t_{new}}{\partial v_i} \alpha) \right] \right| \end{aligned} \quad (25)$$

Figure 1 shows an algorithm for fault diagnosis and location based on above research.

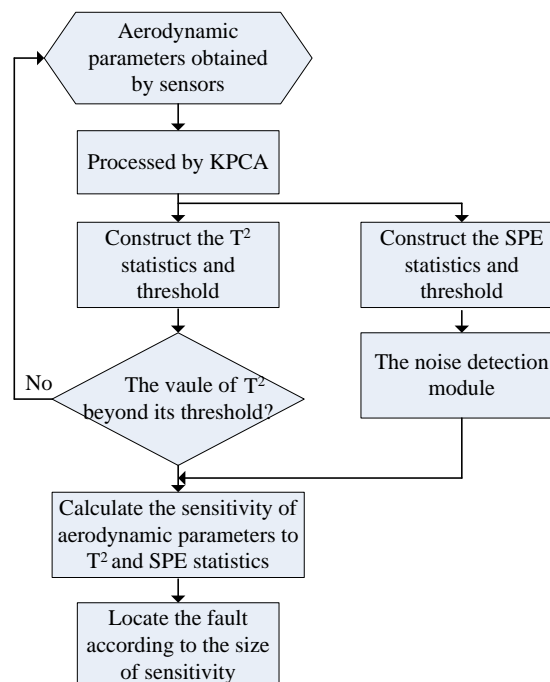


Figure 1. Process of fault diagnosis and location.

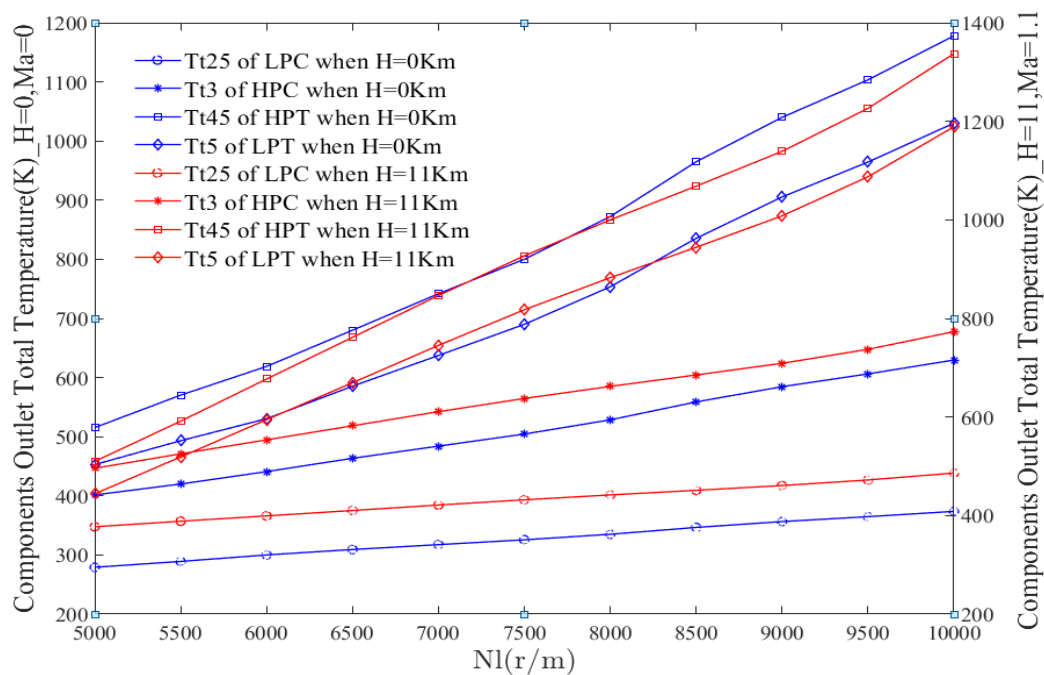
### 3. Results

In order to verify the effectiveness of the method proposed in this paper, certain of twin-spool aviation gas turbine is adopted as the research object. It is widely known that the working condition of the engine is very bad [13,24] (suffering from high pressure, high temperature, high stress, etc.) and the performances of gas turbine components (such as the compressor, rotator, turbine, etc.) are decreasing as working hours increase [25–30]. The initial working parameters of this turbine are shown in Table 1. When these working parameters are determined, the state parameters of the engine are shown in Figures 2 and 3 when the flight altitude and speed are different.

**Table 1.** Initial working parameters of turbine.

Efficiency of LPC	0.868	Pressure ratio of LPC	3.8
Efficiency of HPC	0.878	Pressure ratio of HPC	4.474
Fuel low calorific value	42,900	Total temperature of combustor outlet	1600 K
Efficiency of HPR	0.98	Efficiency of LPR	0.98
Combustion chamber efficiency	0.98	Engine room air entrainment coefficient	0.01
Cooling parameter of HPT	0.03	Efficiency of HPT	0.89
Cooling parameter of LPT	0.01	Efficiency of LPT	0.91
Design speed of LPR	10,000 (r/m)	Design Speed of HPR	16,000 (r/m)

LPC—Low Pressure Compressor; HPC—High Pressure Compressor; HPR—High Pressure Rotor; LPR—Low Pressure Rotor; HPT—High Pressure Turbine; LPT—Low Pressure Turbine.



**Figure 2.** Total temperatures of components.

In the experiment, the measured parameters include the total temperature at the outlets of low pressure compressor LPC (Tt25), total temperature of how pressure compressor HPC (Tt3), total temperature of the high pressure turbine HPT (Tt45), and total temperature of the low pressure turbine LPT (Tt5). In addition, the total pressure at the outlets of the low pressure compressor LPC (Pt25), total pressure of the high pressure compressor HPC (Pt3), total pressure of the high pressure turbine HPT (Pt45), and the total pressure of the low pressure turbine LPT (Pt5) are included. Two faults occurred at the 2600th sampling moment: one is the misalignment of the LPC rotor, and another one is the crack generation of the LPT blade. The proposed method is adopted to detect and locate the faults. Figures 4–7 are the diagrams of fault diagnosis. At the 2600th sampling time, the faults

of the LPC and LPT are generated, respectively. In Figure 2, the value of the  $T^2$  statistic is smooth and lower than its threshold before the 2600th sampling time. Due to the occurrence of fault at the 2600th sample, the curve takes a large jump and exceeds its threshold. In Figure 5, the SPE statistic approaches the threshold at some time before the occurrence of fault. Since the SPE statistic mainly contains noise information, KPCA processing cannot eliminate the noise completely. When the noise amplitude increases, the value of the SPE statistic may exceed its threshold, which has been introduced in Equations (13) and (15). This does not affect the fault diagnosis of the components. Figures 6 and 7 show the fault detection of the LPT and the detection results are similar with those of the LPC. The fault location algorithm mentioned above is used to locate the fault of the LPC, HPC, LPT, and HPT. The location of the results are shown in Table 2.

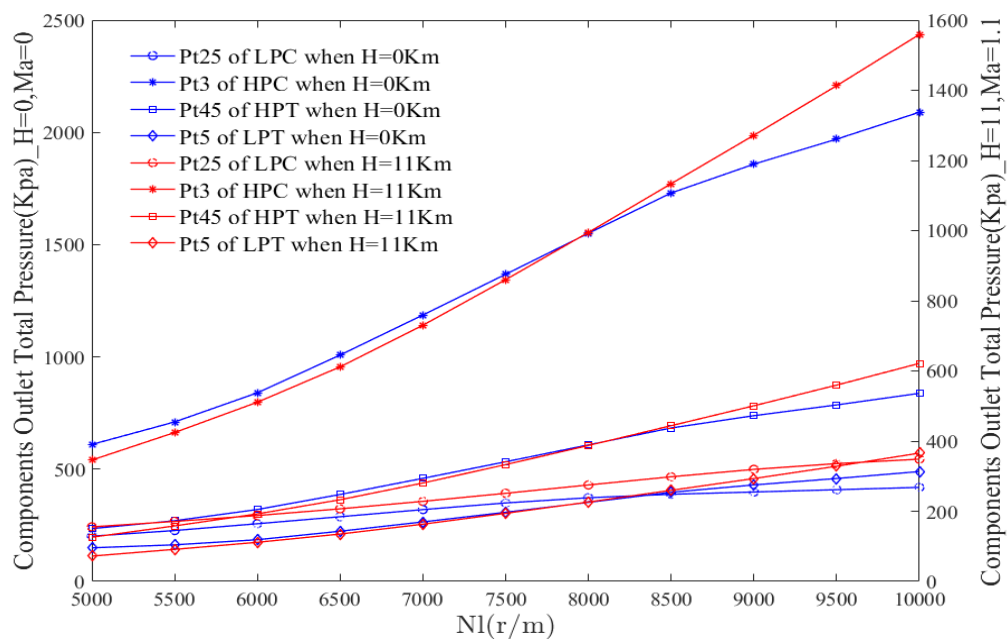


Figure 3. Total pressures of components.

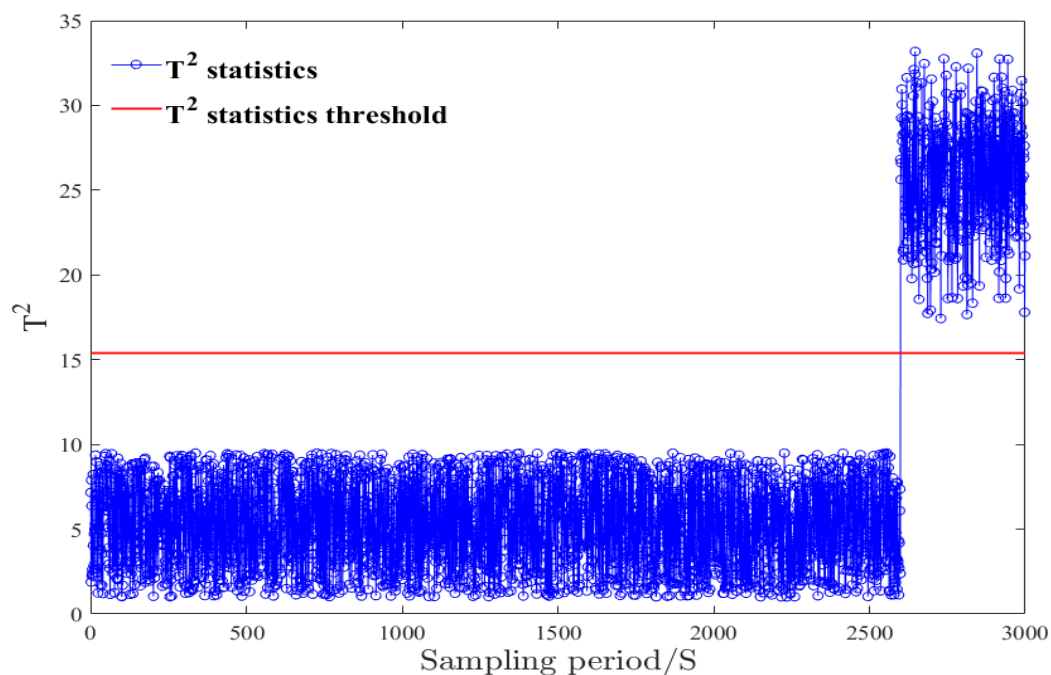


Figure 4. Value of the  $T^2$  statistic under the condition of an LPC (low pressure compressor) fault.



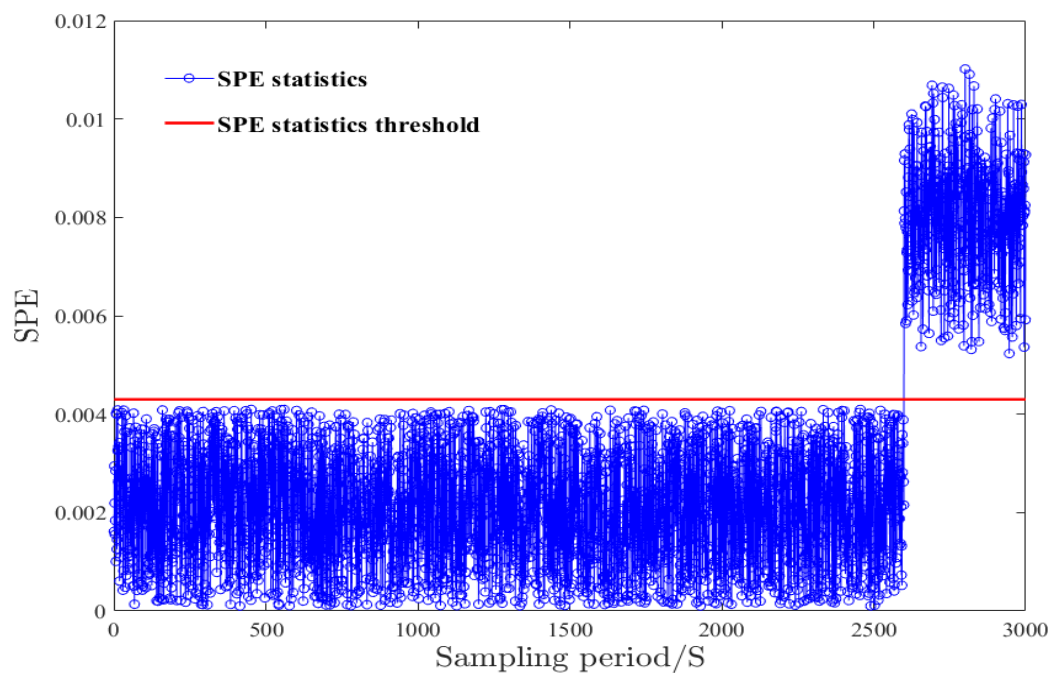


Figure 5. Value of SPE statistic under the condition of LPC fault.

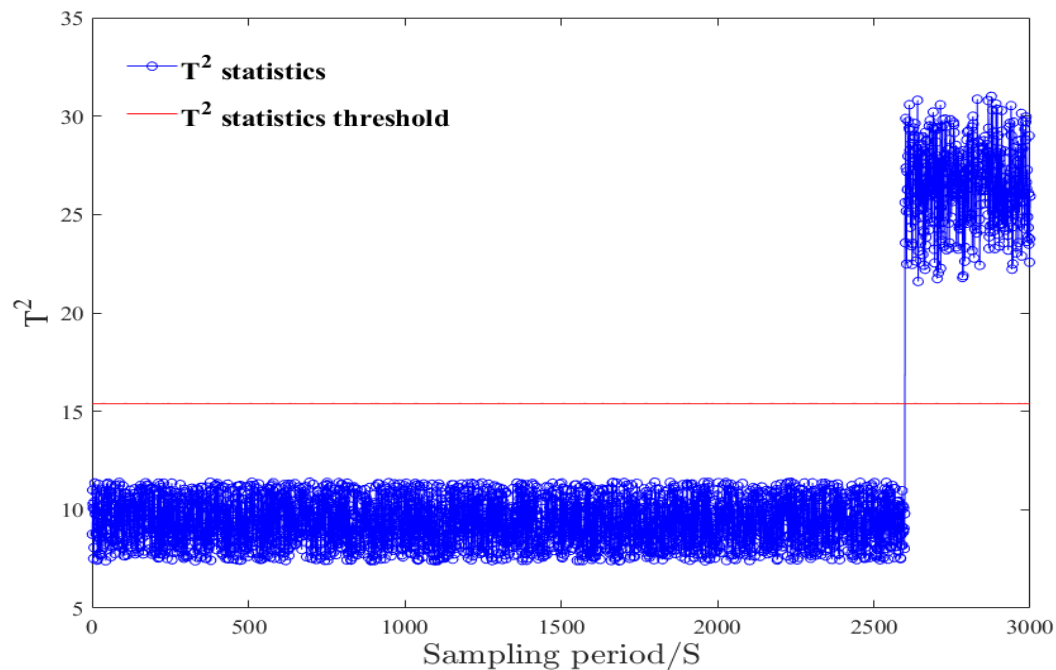
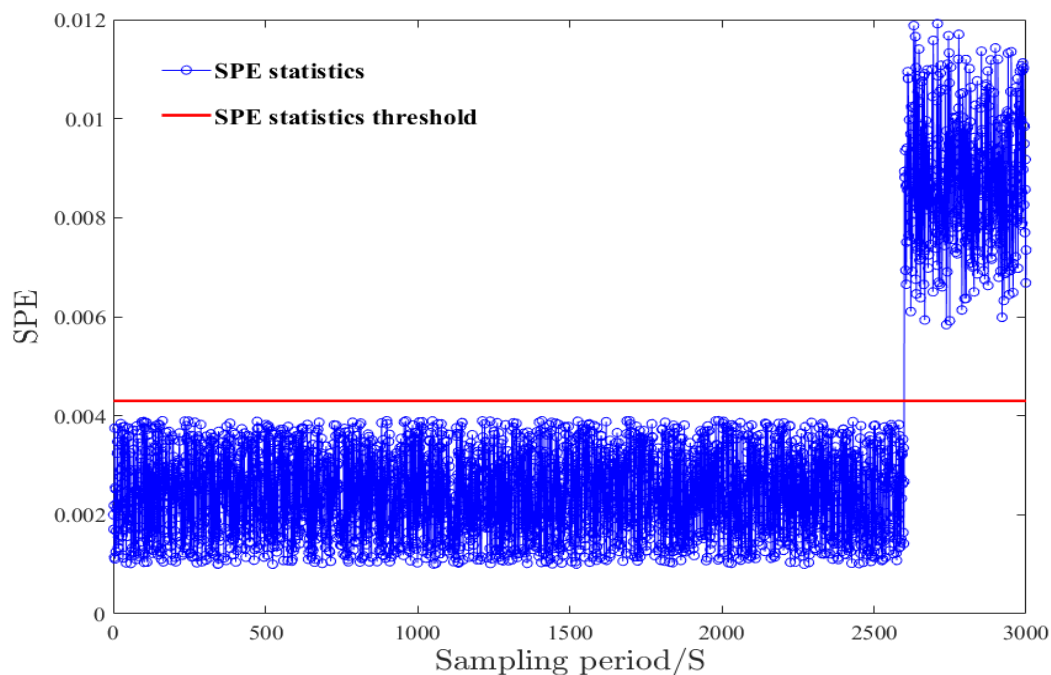


Figure 6. Value of the  $T^2$  statistic under the condition of an LPT fault.

Table 2 shows that when any part of gas turbine components is fault, the sensitivity of measured parameters of faulty part to the statistics of  $T^2$  and SPE is greater than that of normal ones. Take LPC as an example to illustrate the result. If the efficiency coefficient of LPC decreased by 1% due to the misalignment of LPC rotor, the measured parameters at the outlet of the LPC fluctuates firstly. Sensitivities of total temperature and total pressure at the outlet of LPC measured by the sensors to  $T^2$  statistic are 0.3239 and 0.6271, which are obviously higher than those of other measured parameters. The sensitivities to the SPE statistic are both 0.125, which are also higher than those of other parameters. The fault location method can locate the fault to the each component. Figure 8 shows the sensitivity distribution spectrums of the measured parameters when the gas turbine is working.



**Figure 7.** Value of the SPE (Squared Prediction Error) statistic under the condition of an LPT fault.

**Table 2.** The sensitivity of measured parameters.

	Sensitivity	F <sub>LPC</sub>	F <sub>HPC</sub>	F <sub>LPT</sub>	F <sub>HPT</sub>
Sensitivity of measured parameters to statistics	C <sub>T<sup>2</sup>_T25</sub>	0.3239	0.0397	0.037	0.0279
	C <sub>T<sup>2</sup>_T3</sub>	0.0045	0.6043	0.0215	0.0133
	C <sub>T<sup>2</sup>_T45</sub>	0.0144	0.0789	0.0584	0.1764
	C <sub>T<sup>2</sup>_T5</sub>	0.005	0.0248	0.2073	0.0177
	C <sub>T<sup>2</sup>_P25</sub>	0.6271	0.063	0.0554	0.0448
	C <sub>T<sup>2</sup>_P3</sub>	0.0042	0.1014	0.0246	0.0195
	C <sub>T<sup>2</sup>_P45</sub>	0.0156	0.0739	0.0713	0.6866
	C <sub>T<sup>2</sup>_P5</sub>	0.0049	0.014	0.5244	0.0138
	C <sub>SPE_T25</sub>	0.1250	0.1249	0.1249	0.1249
	C <sub>SPE_T3</sub>	0.1249	0.125	0.1249	0.1249
	C <sub>SPE_T45</sub>	0.1249	0.1249	0.1249	0.125
	C <sub>SPE_T5</sub>	0.1249	0.1249	0.125	0.1249
	C <sub>SPE_P25</sub>	0.1250	0.1249	0.1249	0.1249
	C <sub>SPE_P3</sub>	0.1249	0.125	0.1249	0.1249
	C <sub>SPE_P45</sub>	0.1249	0.1249	0.1249	0.125
	C <sub>SPE_P5</sub>	0.1249	0.1249	0.1245	0.1249

It can be seen that before the fault occurs in this figure, the sensitivity distribution curves are gentle and the differences between the sensitivity curves are not obvious. When a low pressure compressor failure occurs, the curves representing the sensitivity of the LPC increased sharply in a short time, and the values are significantly higher than others. In addition, according to the rule of failure caused by the degradation of gas turbine components, if the performance of a component degrades to a certain extent and is about to fail, the degradation speed will be accelerated until the failure occurs. In this process, the measured parameters at the outlet of deteriorating components will deviate from the real value as the deterioration of performance. In the sensitivity distribution spectrum, the sensitivity of the measured values of the deteriorating components will increase continuously, and the fault prediction can be realized by comparing the changes in sensitivity.

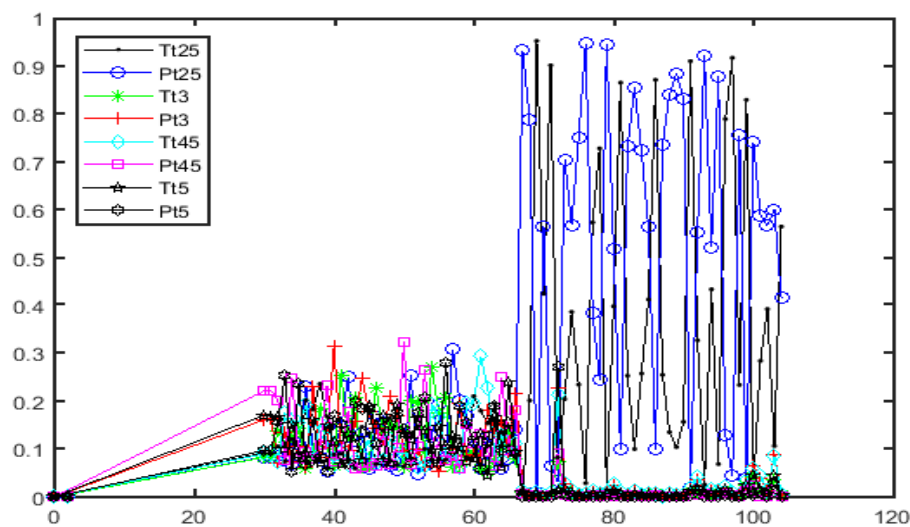


Figure 8. Sensitivity distribution spectra of the tested parameters.

In addition, due to the influence of harsh working circumstance, the sensor outputs may seriously deviate from their actual values and this may lead to the misdiagnosis. It is essential to keep output within a reasonable range. According to the working principle of gas turbines (taking an aero gas engine as an example), the power and flow balance conditions must be observed when the turbine is under normal working conditions and all parameters remain unchanged or fluctuate in a small range. If the state of turbine changes due to the variation of control parameters, all the aerodynamic parameters will bound to change greatly, reflecting an anomaly of sensor measurements. Another case is that if only a few measurements are abnormal, according to the working principle of gas turbines and the balance conditions, it can be known that the anomalies are caused by noise or the fault of the sensors and the measurements must be restored. The process to detect the abnormal value and restore the measurements is shown in Figure 9.

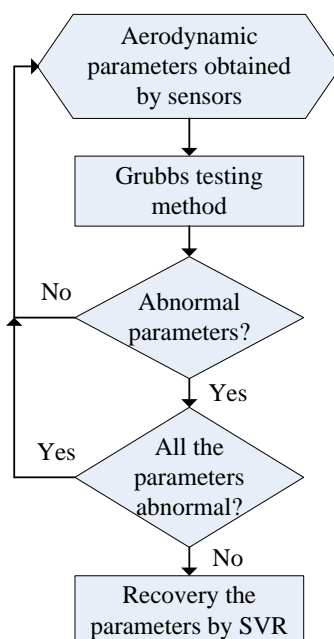


Figure 9. Processing steps for outliers.

Firstly, Grubb's method for testing is adopted to check if the parameters of sensors are abnormal or not. It is important to note that the value of detection level  $\alpha$  must be determined according to the

variation of the aerodynamic parameters. If the checked parameters are abnormal, all the parameters should be tested by Grubb's method. Then count the number of parameters which are abnormal. If the number is less than the quantity of aerodynamic parameters for fault diagnosis, it indicates that the sensors are faulty or the influence of noise increases. In such a situation, the abnormal parameters must be restored by a support vector machine (SVR) [31] to ensure the fault location proceeds smoothly.

#### 4. Conclusions

In this paper, a novel method to locate the fault of a gas turbine is proposed. Kernel principal component analysis is adopted to detect the occurrence of fault. Based on the analysis of the fault indicator and aerodynamic parameters, the partial derivative of the  $T^2$  and SPE statistics to aerodynamic parameters are calculated. The results are used to represent the influence degree of fault to these parameters and the fault location can be realized by different influence degrees. There are four conclusions can be drawn, as follows:

(1) The KPCA is an effective way to detect the fault of a gas turbine.  $T^2$  and SPE statistics, and their corresponding thresholds must be constructed. By comparing the size of the  $T^2$  statistic and its threshold, the on-line fault diagnosis of a gas turbine can be realized. Furthermore, both statistics help to locate the fault.

(2) The fault location of gas turbine is realized by calculate the partial derivatives of  $T^2$  to aerodynamic parameters. The size of partial derivatives represent the sensitivity degrees of aerodynamic parameters to fault. Based on the balance working condition of gas turbine, fault location can be achieved according to the size of partial derivatives.

(3) Sensitivity distribution spectra can be used to represent the performance degradation of the gas turbine and identify a potential fault. When a fault which resulted by performance degradation is about to occur, the partial derivatives of the aerodynamic parameters associated with this fault will change dramatically and this change is easily reflected by the sensitivity distribution spectra.

(4) The method proposed in this paper is used to locate the fault of a gas turbine and how to recognize the fault is not discussed. Whether this method can be used in fault identification needs to be verified in the follow-up work.

**Author Contributions:** Data curation: L.Z.; funding acquisition: W.L.; methodology: L.Z.; project administration: W.L. and Y.L.; supervision: Y.L.; writing—original draft, L.Z.; writing—review and editing: L.Z.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 51875371 and National Green Manufacturing System Plan Number: [2017] 327.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Siddiqui, M.E.; Almitani, K.H. Energy Analysis of the S-CO<sub>2</sub> Brayton Cycle with Improved Heat Regeneration. *Processes* **2019**, *7*, 3. [\[CrossRef\]](#)
2. Lee, H.G.; Huh, J.H. A Cost-Effective Redundant Digital Excitation Control System and Test Bed Experiment for Safe Power Supply for Process Industry 4.0. *Sensors* **2018**, *7*, 85. [\[CrossRef\]](#)
3. Lu, F.; Ju, H.F.; Huang, J.Q. An improved extended Kalman filter with inequality constraints for aero engine health monitoring. *Aerosp. Sci. Technol.* **2016**, *58*, 36–47. [\[CrossRef\]](#)
4. Zeng, L.; Long, W.; Li, Y.Y. Performance Degradation Diagnosis of Gas Turbine Based on Improved FUKF. *J. Southwest Univ.* **2018**, *53*, 873–878. [\[CrossRef\]](#)
5. Gao, Z.W.; Ding, S.X.; Cecati, C. Real-time fault diagnosis and Fault-tolerant Control. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3752–3756. [\[CrossRef\]](#)
6. Gao, Z.W.; Sheng, S.W. Real-time monitoring, prognosis, and resilient control for wind turbine systems. *IEEE Trans. Ind. Electron.* **2018**, *116*, 1–4. [\[CrossRef\]](#)
7. Simani, S.; Farsoni, S.; Castaldi, P. Fault Diagnosis of a Wind Turbine Benchmark via Identified Fuzzy Models. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3775–3782. [\[CrossRef\]](#)

8. Sanchez, S.; Escobet, T.; Puig, V.; Odgaard, P.F. Fault Diagnosis of an Advanced Wind Turbine Benchmark Using Interval-Based ARRs and Observers. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3783–3792. [[CrossRef](#)]
9. Gao, Z.; Cecati, C.; Ding, S.X. A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3757–3767. [[CrossRef](#)]
10. Gao, Z.; Cecati, C.; Ding, S.X. A Survey of Fault Diagnosis and Fault-tolerant Techniques—Part II: Fault diagnosis with knowledge-based and hybrid/active approaches. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3768–3774. [[CrossRef](#)]
11. Zhang, Y.; Bingham, C.M.; Gallimore, M. Fault Detection and Diagnosis Based on Extensions of PCA. *Adv. Mil. Technol.* **2013**, *8*, 1–15. [[CrossRef](#)]
12. Zhu, J.L.; Ge, Z.Q.; Song, Z.H. HMM-Driven Robust Probabilistic Principal Component Analyzer for Dynamic Process Fault Classification. *IEEE Trans. Ind. Electron.* **2015**, *62*, 3814–3821. [[CrossRef](#)]
13. Zhang, P. Aero Engine Fault Diagnostics Based on Kalman Filter. Ph.D. Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2008.
14. Vasile, P.; Ron, J.P.; Faisel, J.U.; Joseba, Q. Fault Diagnosis of An Industrial Gas Turbine Using Neuro-fuzzy Methods. *World Congr.* **2002**, *35*, 1630–1636. [[CrossRef](#)]
15. Che, C.C.; Wang, H.W.; Ni, X.M.; Hong, J.Y. Fault Fusion Diagnosis of Aero-engine Based on Deep Learning. *J. Beijing Univ. Aeronaut. Astronaut.* **2018**, *44*, 621–628. [[CrossRef](#)]
16. Tayarani-Bathaie, S.S.; Khorasani, K. Fault Detection and Isolation of Aero Engines Using a Bank of Neural Networks. *Process Control* **2015**, *36*, 22–41. [[CrossRef](#)]
17. Lei, W.; Zhang, J.X.; Xie, S.S.; Ren, L.T.; Hu, J.H. Fault Estimation for Actuators in Aero engine Control System Based on Fault Decomposition. *Fire Control Command Control* **2014**, *39*, 127–130.
18. Fuat, D.; Jose, A.R.; Ahmet, P. A strategy for Detection and Isolation of Sensor Failures and Process Upsets. *Chemom. Intell. Lab. Syst.* **2001**, *55*, 109–123.
19. Wold, S. Exponentially weighted moving principal component analysis and projection to latent structures. *Chemom. Intell. Lab. Syst.* **1994**, *21*, 149–161. [[CrossRef](#)]
20. Navi, M.; Meskin, N.; Davoodi, M. Sensor Fault Detection and Isolation of An Industrial Gas Turbine Using Partial Adaptive KPCA. *J. Process Control* **2015**, *2*, 37–48. [[CrossRef](#)]
21. Chen, S.G.; Wu, X.J.; Yin, H.F. KPCA Method Based on Within-class Auxiliary Training Samples and Its Application to Pattern Classification. *Pattern Anal. Appl.* **2016**, *3*, 749–767. [[CrossRef](#)]
22. Alain, R. Variable Selection Using SVM-based Criteria. *J. Mach. Learn. Res.* **2003**, *3*, 1357–1370. [[CrossRef](#)]
23. Jiang, W.L.; Wu, S.X.; Liu, S.Y. Exponentially Weighted Dynamic Kernel Principal Component Analysis Algorithm and Its Application in Fault Diagnosis. *J. Mech. Eng.* **2011**, *47*, 63–67. [[CrossRef](#)]
24. Zhang, K.Y.; Li, Z.G.; Li, J. Effect of gas turbine endwall misalignment with slashface leakage on the blade endwall aerothermal performance. *Int. J. Heat Mass Transf.* **2018**, *131*, 1247–1259. [[CrossRef](#)]
25. Dewallef, P.; Romessis, C.; Leonard, O.; Mathioudakis, K. Combining classification techniques with Kalman filters for aircraft engine diagnostics. *J. Eng. Gas Turbines Power* **2006**, *128*, 595–603. [[CrossRef](#)]
26. Huang, M.Y.; Dey, S. Stability of Kalman filtering with Markovian packet losses. *Automatica* **2007**, *43*, 598–607. [[CrossRef](#)]
27. Chen, Y. Gas Path Fault Diagnosis for Turbojet Engine Based on Nonlinear Model. Ph.D. Thesis, Nanjing University of Aeronautics and Astronautics, Nanjing, China, 2014.
28. Zhang, P.; Huang, J.Q. SRUKF Research on Aero engine for Gas Path Component Fault Diagnosis. *J. Aeros. Power* **2008**, *23*, 169–172.
29. Kurzke, J. Gas turbine cycle design methodology: A comparison of parameter variation with numerical optimization. *J. Eng. Gas Turbines Power* **1999**, *121*, 6–11. [[CrossRef](#)]
30. Lu, F.; Huang, J.Q. Contracted Kalman Filter Estimation for Turbofan Engine Gas-path Health. *Control Theory Appl.* **2012**, *29*, 1543–1550. [[CrossRef](#)]
31. Zeng, L.; Long, W. Prediction Method of Metal Content of Aero Engine Lubricating Oil Based on SVR. *J. Sichuan Univ.* **2016**, *48*, 161–164. [[CrossRef](#)]

